Measure valued differentiation for Markov Chains

Heidergott, B.F.; Vazquez-Abad, F.

published in
Journal of Optimization Theory and Applications
2008

DOI (link to publisher)
10.1007/s10957-007-9297-7

document version
Publisher's PDF, also known as Version of record

J Optim Theory Appl (2008) 136: 187–209
DOI 10.1007/s10957-007-9297-7

Measure-Valued Differentiation for Markov Chains

B. Heidergott · F.J. Vázquez-Abad

Published online: 30 October 2007

Abstract This paper addresses the problem of sensitivity analysis for finite-horizon performance measures of general Markov chains. We derive closed-form expressions and associated unbiased gradient estimators for the derivatives of finite products of Markov kernels by measure-valued differentiation (MVD). In the MVD setting, the derivatives of Markov kernels, called $\mathcal{D}$-derivatives, are defined with respect to a class of performance functions $\mathcal{D}$ such that, for any performance measure $g \in \mathcal{D}$, the derivative of the integral of g with respect to the one-step transition probability of the Markov chain exists. The MVD approach (i) yields results that can be applied to performance functions out of a predefined class, (ii) allows for a product rule of differentiation, that is, analyzing the derivative of the transition kernel immediately yields finite-horizon results, (iii) provides an operator language approach to the differentiation of Markov chains and (iv) clearly identifies the trade-off between the generality of the performance classes that can be analyzed and the generality of the classes of measures (Markov kernels). The $\mathcal{D}$-derivative of a measure can be interpreted in terms of various (unbiased) gradient estimators and the product rule for $\mathcal{D}$-differentiation yields a product-rule for various gradient estimators.

Communicated by Yu Chi Ho.

Part of this work was done while the first author was with EURANDOM, Eindhoven, Netherlands, where he was supported by Deutsche Forschungsgemeinschaft under Grant He3139/1-1. The work of the second author was partially supported by NSERC and FCAR grants of the Government of Canada and Québec.

B. Heidergott (✉)
Vrije Universiteit and Tinbergen Institute, Department of Econometrics and Operations Research, Amsterdam, Netherlands
e-mail: bheidergott@feweb.vu.nl

F.J. Vázquez-Abad
Department of Mathematics and Statistics, and ARC Special Research Centre for Ultra-Broadband Information Networks, University of Melbourne, Melbourne, Australia

Keywords Gradient estimation · Simulation · Perturbation analysis · Measure-valued differentiation

1 Introduction

Many real-world systems in manufacturing, transportation, communication networks, or finance can be modeled by general state-space Markov chains, such as generalized semi-Markov processes (see [1]). The past two decades have witnessed an increased attention to the study of gradient estimation for discrete event driven systems (see [1–3]), with the aim of finding better and more efficient control methods.

The motivation for the present paper is to establish a mathematical framework that comprises (most of) the existing gradient estimation methods for Markov chains. Our approach summarizes the proof techniques and ideas that are known in the literature in order to establish a unified theory of gradient estimation. In our view, such a framework has to (i) provide a general and meaningful concept of differentiation, (ii) satisfy a product rule of differentiation for this concept of differentiation, (iii) allow statements obtained within this theory to be translated into unbiased gradient estimators, and finally (iv) deal with random horizon problems.

In this paper, we show that measure-valued differentiation (MVD) provides the means to establish a unified theory of gradient estimation. In particular, we address here topics (i) to (iii). The fact the MVD is an operator language approach will prove most helpful when going from (i) to (ii). Topic (iv) can be found in [4].

The paper is organized as follows. Section 2 discusses various approaches to the gradient estimation problem. We illustrate to what extent these methods already have features of the intended theory. In Sect. 3, we introduce measure-valued derivatives and we establish the key technical result, which is the product rule of measure-valued differentiation. In Sect. 4, we show how the conditions of the product rule can be verified in various scenarios that are of importance in applications. For example, when only bounded functions are considered, the conditions for the product rule can be expressed in a very simple manner. Section 5 shows how the expressions produced by the product rule (and containing signed measures) can be turned into various types of gradient estimators, such as those typically obtained from SPA, the Score function method or weak derivatives.

All the proofs and detailed developments of our claims are provided in a web supplement [5], which will not be cited further to avoid unnecessary repetitions.

2 Background and Motivation

Let $\{X_\theta(n)\}$, with $\theta \in \Theta \subset \mathbb{R}$, be a Markov Chain with (arbitrary) state space S defined on a common probability space $(\Omega, \mathfrak{F}, \mathbb{P})$, where Θ is the set of (control) parameters such that $(\Omega, \mathfrak{F}, \mathbb{P})$ is independent of θ and $X_\theta(n)$ is well defined on Θ. The problem of sensitivity analysis can be phrased as follows. For performance functions $g : S \to \mathbb{R}$, find conditions such that, for $n \in \mathbb{N}$,

$$\frac{d}{d\theta}\mathbb{E}\big[g(X_\theta(n), \ldots, X_\theta(1))\big] \tag{1}$$

exists and can be obtained in a closed form expression. When derivatives are defined, it is sufficient that Θ be a neighborhood of the point θ of interest. The last two decades have witnessed a great interest in the problem of finding unbiased estimators for the expression in (1), called gradient estimators, see for example [1–3] and [6–8]. The term *sensitivity analysis* is often used to refer to this problem. The methods available are legion and even experts find it difficult to oversee them all. However, the following three major approaches can be identified: *smoothed perturbation analysis* (SPA), *score function* and *weak derivatives*, which will be described in what follows in more detail.

2.1 Smoothed Perturbation Analysis

In the sample-path analysis setting, the dependency of the expectation in the parameter θ is expressed entirely through the performance. If the sample performance is almost surely Lipschitz continuous in θ, then the sample path derivative $dg(X_\theta(n,\omega))/d\theta$ is unbiased for the gradient (1), yielding the so-called infinitesimal perturbation analysis, see [1]. In the presence of discontinuities, conditioning can be used to integrate (or "smooth out") such discontinuities, see [2]. Two approaches can be identified. Let $E_{\theta,x}(k)$ denote the event $\{X_\theta(k)=x\}$. The first seeks an estimator of the form

$$\frac{d}{d\theta}\mathbb{E}[g(X_\theta(k+1))|E_{\theta,x}(k)] = \mathbb{E}\left[\frac{d}{d\theta}\mathbb{E}[g(X_\theta(k+1))\,|\mathcal{G}]\,\middle|\,E_{\theta,x}(k)\right],$$

where $\mathcal{G}$ is a smoothing σ-field, and $H(\theta)=\mathbb{E}[g(X_\theta(k+1))|E_{\theta,x}(k),\mathcal{G}]$ is a.s. Lipschitz continuous. It is often very difficult to identify such conditioning fields in practice.

The second approach [2] prescribes an analysis of a perturbed path using the same trajectory ω for θ and for $\theta\pm\Delta$ and conditioning on the (rare) events where discontinuities may occur. Under this formulation the *nominal* and *perturbed* processes $X_\theta, X_{\theta+\Delta\theta}$ share a common filtration. Let $\{X_\theta(k,x), k\in\mathbb{N}\}$ the process started at $X_\theta(0)=x$ and consider evaluating the sensitivities of the one-step expectation. By the Markov property, this sensitivity is

$$\frac{d}{d\theta}\mathbb{E}[g(X_\theta(k+1))|E_{\theta_0,x}(k)]\Big]_{\theta=\theta_0} = \lim_{\Delta\theta\to 0}\mathbb{E}\left[\frac{g(X_{\theta_0+\Delta\theta}(1,x))-g(X_{\theta_0}(1,x))}{\Delta\theta}\right].$$

The random variable $H_x(\theta)=g(X_\theta(1,x))$ may fail to be a.s. Lipschitz continuous, but it is possible to divide the state space into a set $\mathcal{A}_x^*(\Delta\theta,\theta)$ containing only trajectories where $H_x(\theta)$ is Lipschitz continuous, and the so-called *critical set* $\mathcal{A}_x(\Delta\theta,\theta)=\{\omega : |g(X_{\theta+\Delta\theta}(1,x;\omega))-g(X_\theta(1,x;\omega))|>\alpha\Delta\theta\}$, for some $\alpha>0$. It is assumed here that for each state x the limit $\mathcal{A}_x(\Delta\theta,\theta)\to\mathcal{A}_x(\theta)$ exists, for some measurable set $\mathcal{A}_x(\theta)$ and that the limit (called *critical rate*) $\lim_{\Delta\theta\to 0}(1/\Delta)\mathbb{P}(\mathcal{A}_x(\Delta\theta,\theta))=p'_\theta(x)>0$ exists and is finite. This implies that for each $x\in S$, $\lim_{\Delta\theta\to 0}\mathbb{P}(\mathcal{A}^*(\Delta\theta,\theta))=1$. In addition, if the discontinuity itself is absolutely integrable, that is: $\mathbb{E}[|g(X_{\theta+\Delta\theta}(1,x))-g(X_\theta(1,x))||\mathcal{A}_x(\Delta\theta,\theta)]<\infty$, then the dominated convergence theorem yields

$$\lim_{\Delta\theta \to 0} \mathbb{E}\left[\frac{g(X_{\theta+\Delta\theta}(1,x)) - g(X_\theta(1,x))}{\Delta\theta}\right]$$
$$= \mathbb{E}\left[\frac{d}{d\theta} H_x(\theta)\right] + \mathbb{E}[g(X_{\theta^+}(1,x)) - g(X_\theta(1,x)) | \mathcal{A}_x(\theta)] p'_\theta(x),$$

where $X_{\theta^+}(1,x)$ denotes the limit of $X_{\theta+\Delta\theta}(1,x)$ as $\Delta\theta \downarrow 0$. The term inside the first expectation is known as the *IPA term* and the second, as the *SPA term* of the derivative estimator. The effect of conditioning on the so-called critical events is to partially integrate the discontinuities via the critical rate p'_θ.

In the foregoing, only one transition was affected by the perturbation of θ. When studying the process $\{X_\theta(n)\}$, the perturbations affect the entire trajectories. As done in [2], the expectation is rewritten in terms of filtered Monte Carlo, conditioning on each step. Under the assumed integrability conditions, the overall effect of the SPA term is obtained as if only one-step transitions were perturbed at a time, that is (assuming no IPA contribution),

$$\frac{d}{d\theta}\mathbb{E}[g(X_\theta(k+1))]$$
$$= \mathbb{E}\left[\sum_{i=0}^{k} \mathbb{E}[g(X_\theta^{(i)}(k+1)) - g(X_\theta(k+1)) | \mathcal{A}_{X_\theta(i)}(\theta)] p'_\theta(X_\theta(i))\right], \tag{2}$$

where the process $\{X_\theta^{(i)}(n)\}$ is the limiting process from a perturbation $\theta + \Delta\theta$ *at the ith transition only*. To show the validity of the expression above and to obtain the sample path estimators, the crucial step when using the path-wise analysis is to show that for small changes in θ, the discontinuous effect of the perturbation of the whole trajectory is only *local*: discontinuities initiate at each transition and then propagate. To summarize, SPA involves a careful path-wise analysis of the propagation of delays and their effect on a given performance measure. While SPA offers great flexibility, proofs of unbiasedness are often very cumbersome, because the effect of a perturbation on the *entire* sample path has to be studied. Furthermore, the results for SPA only hold for individual performance functions, so different performance functions require a entirely new proofs.

Under our interpretation in terms of MVD, the term p'_θ in (2) represents the derivative of a probability distribution, and calculating the overall gradient in (1) corresponds to applying a product rule of differentiation to a product measure. In this paper we derive such a product rule of measure-valued differentiation, which can be applied to SPA. More precisely, our product rule for measure-valued differentiation provides the sensitivity of the entire sample path out of a local analysis. Put another way, the analysis of the propagation of delays is taken care of by the product rule.

In contrast to SPA, MVD does not require to find a smoothing σ-field. While finding a $\mathcal{D}$-derivative can be as time consuming as finding the appropriate smoothing σ-field in SPA, the advantage of MVD is that one can always come up with a Hahn-Jordan decomposition as $\mathcal{D}$-derivative (refer to Sect. 3). In applications, it is desirable to use a $\mathcal{D}$-derivative that has a nice interpretation and that can be easily implemented. It is at this point that time and effort are required to come up with a $\mathcal{D}$-derivative tailored to the problem. Here 'tailored' means that the $\mathcal{D}$-derivative has to be adapted to the random dynamic of the system, whereas the actual performance measure is of no concern (which is in contrast to SPA).

2.2 Weak Derivatives

In this section, we review briefly the concept of weak differentiation of probability measures as introduced by Pflug [8]. Let $(S,\mathcal{S})$ denote a Polish measurable space.[1] For most applications, $S \subset \mathbb{R}^d$ and $\mathcal{S}$ represents the σ-field of events that are Borel subsets of $\mathbb{R}^d$. Let $\mathcal{M}=\mathcal{M}(S,\mathcal{S})$ denote the set of finite signed measures on $(S,\mathcal{S})$, and $\mathcal{M}_1=\mathcal{M}_1(S,\mathcal{S})\subset\mathcal{M}$ the set of probability measures. Denote by $C_b := C_b(S)$ the set of bounded, continuous mappings $g: S\to\mathbb{R}$. For any signed measure ν on $(S,\mathcal{S})$ there exists a set $G\in\mathcal{S}$, such that $[\nu]^+(A) := \nu(A\cap G)\geq 0$ and $[\nu]^-(A) := -\nu(A\cap G^c)\geq 0$ for any $A\in\mathcal{S}$. In particular, the set G is implicitly defined via

$$\nu(G)=\sup\{A\in\mathcal{S}:\nu(A)\}. \tag{3}$$

The measures $[\nu^+]$ and $[\nu^-]$ are positive measures on $(S,\mathcal{S})$ and the pair $([\nu]^+,[\nu]^-)$ is called the *Hahn-Jordan decomposition* of ν. The Hahn-Jordan decomposition is unique in the sense that if $\hat{G}$ is another set, such that $\nu(A\cap\hat{G})\geq 0$ and $\nu(A\cap\hat{G}^c)\leq 0$ for any $A\in\mathcal{S}$, then $\nu(A\cap G)=\nu(A\cap\hat{G})$ for any $A\in\mathcal{S}$. A signed measure is called *finite* if $[\nu]^+$ and $[\nu]^-$ are finite measures. Integration with respect to a signed measure is defined through

$$\int_S g(s)\nu(ds)=\int_S g(s)[\nu]^+(ds)-\int_S g(s)[\nu]^-(ds),$$

provided that the terms on the right-hand side are finite.

Definition 2.1 A measure $\mu_\theta\in\mathcal{M}_1$ is called *weakly differentiable at* θ if a signed finite measure $\mu'_\theta\in\mathcal{M}$ exists, such that, for all $g\in C_b$, it holds that

$$\lim_{\Delta\to 0}\frac{1}{\Delta}\left(\int_S g(s)\mu_{\theta+\Delta}(ds)-\int_S g(s)\mu_\theta(ds)\right)=\int_S g(s)\mu'_\theta(ds).$$

Note that

$$\mu'_\theta(S)=\int_S \mu'_\theta(ds)=0$$

(take g=1), so that μ'_θ can be written as difference between two probability measures (apply, for example, the Hahn-Jordan decomposition).

[1] A topological space is called *separable* if it contains a countable dense set. It is called *Polish* if there exists a metric compatible with the topology under which the space is complete and separable; see e.g. [9].

Definition 2.2 A triple $(c_\theta, \mu_\theta^+, \mu_\theta^-)$ is called a *weak derivative* of μ_θ, where $\mu_\theta^\pm \in \mathcal{M}_1$, if for all continuous bounded functions $g \in C_b$ it holds that

$$\begin{aligned}\int_S g(s)\mu_\theta'(ds) &= \lim_{\Delta\to 0}\frac{1}{\Delta}\left(\int_S g(s)\mu_{\theta+\Delta}(ds) - \int_S g(s)\mu_\theta(ds)\right)\\ &= c_\theta\left(\int_S g(s)\mu_\theta^+(ds) - \int_S g(s)\mu_\theta^-(ds)\right).\end{aligned} \tag{4}$$

The probability measure μ_θ^+ is called the (normalized) positive part of μ_θ' and μ_θ^- is called the (normalized) negative part of μ_θ', respectively. Note that the weak derivative is not unique. We illustrate this with the following example.

Example 2.1 Let $S = [0, \infty)$ and η_θ the exponential distribution with mean θ. Let $f_\theta(x) = \theta\exp(-\theta x)$ denote the Lebesgue density of η_θ; then it holds, for any $g \in C_b$,

$$\begin{aligned}\frac{d}{d\theta}\int g(x)\eta_\theta(dx) = \frac{d}{d\theta}\int g(x) f_\theta(x)dx &= \int g(x)\frac{d}{d\theta} f_\theta(x)dx\\ &= \int g(x)(1-\theta x)e^{-\theta x}dx\\ &= \frac{1}{\theta}\left(\int g(x) f_\theta(x)dx - \int g(x)h_\theta(x)dx\right),\end{aligned}$$

where h_θ is the density of the gamma distribution $\Gamma(2, \theta)$. Hence, η_θ is weakly differentiable and an instance of a weak derivative of μ_θ is given by $(1/\theta, \eta_\theta, \Gamma(2, \theta))$. On the other hand, the Hahn-Jordan decomposition leads to the representation $((\theta e)^{-1}, \mu_\theta^+, \mu_\theta^-)$ with

$$\mu_\theta^+(A) = \int_0^{1/\theta} 1_A(x)(\theta - \theta^2 x)e^{1-\theta x}dx$$

and

$$\mu_\theta^-(A) = \int_{1/\theta}^{\infty} 1_A(x)(\theta^2 x - \theta)e^{1-\theta x}dx,$$

for any measurable set A.

From Definition 2.2 it is clear that weak derivatives yield results which hold on C_b. A product rule of weak differentiation for products of independent measures appeared in [8]. However, whether there also exists a product rule of weak differentiation for conditional measures, like Markov kernels, is still an open question. The MVD approach that we introduce in Sect. 3 extends the results of weak derivatives in two aspects: the product form is now established for the product of Markov kernels, and admissible performance functions are more general, no longer requiring (piece-wise) continuity and boundedness.

2.3 Score Function

Assume that $\nu \in \mathcal{M}$ exists, such that μ_θ is absolutely continuous with respect to ν for all $\theta \in \Theta$ and denote the ν-density of μ_θ by f_θ. If f_θ is ν almost surely differentiable with respect to θ and $\int_S \sup_{\theta\in\Theta} |df_\theta(u)/d\theta|\, \nu(du) < \infty$, then for any $g \in C_b$,

$$\frac{d}{d\theta}\int g(u)\mu_\theta(du) = \int g(u)\frac{d}{d\theta} f_\theta(u)\nu(du) = \int g(u)\frac{d}{d\theta}\ln(f_\theta(u))\mu_\theta(du). \quad (5)$$

The mapping $d\ln(f_\theta(u))/d\theta$ is called *score function*.

The score function approach works on C_b [3]. Furthermore, standard calculus implies a product rule for the score function. The key condition for the above approach is that $f_\theta(u) = 0$ implies $df_\theta(u)/d\theta = 0$. In other words, the measureμ'_θ given by

$$\mu'_\theta(B) = \int_B df_\theta(u)/d\theta\,\nu(du), \quad \text{for } B \in \mathcal{S},$$

is absolutely continuous w.r.t. ν and μ_θ. As we will illustrate in the web supplement this restricts the applicability of the score function approach. In addition, the score function estimates suffer typically from variance problems.

Remark 2.1 Note that if (5) holds for any $g \in C_b$, then μ_θ is weakly differentiable. However, the converse is not true. For a counterexample and details, see [5]. To see this, the key observation is that the above score function approach requires that μ_θ as well as μ'_θ are absolutely continuous with respect to the *same* measure ν, which is not required for the measure-valued concept of differentiation.

3 Measure-Valued Differentiation

In this section, we formally present the concept of measure-valued differentiation (MVD), inspired by the concept of weak differentiation, but as we will soon establish, our methodology does not rely on weak topology only. The main result of this section is the proof of the product rule for MVD. Furthermore, using a conditioning approach, we show how the measure-valued differentiability of a Markov kernel can be deduced from that of more elementary distributions.

3.1 Basic Concepts and Definitions

Let $(S_2, \mathcal{S}_2)$ and $(S_1, \mathcal{S}_1)$ be Polish measurable spaces. Recall that $\mathcal{M}(S_2, \mathcal{S}_2)$ denotes the set of finite (signed) measures on $(S_2, \mathcal{S}_2)$ and $\mathcal{M}_1(S_2, \mathcal{S}_2)$ that of probability measures on $(S_2, \mathcal{S}_2)$.

Definition 3.1 The mapping $P : \mathcal{S}_2 \times S_1 \to [0, 1]$ is called a *transition kernel* on $(\mathcal{S}_2, S_1)$ if:

(a) $P(\cdot; s) \in \mathcal{M}(S_2, \mathcal{S}_2)$, for all $s \in S_1$;
(b) $P(B; \cdot)$ is $\mathcal{S}_1$ measurable for all $B \in \mathcal{S}_2$.

P is called a *Markov kernel* on $(\mathcal{S}_2, S_1)$ when $\mathcal{M}(S_2, \mathcal{S}_2)$ can be replaced by $\mathcal{M}_1(S_2, \mathcal{S}_2)$ in (a).

Denote the set of transition kernels on $(\mathcal{S}_2, S_1)$ by $\mathcal{K}(\mathcal{S}_2, S_1)$ and the set of Markov kernels on $(\mathcal{S}_2, S_1)$ by $\mathcal{K}_1(\mathcal{S}_2, S_1)$. If $(S_2, \mathcal{S}_2) \neq (S_1, \mathcal{S}_1)$, then the transition (respectively, Markov) kernel is called *in-homogeneous*, whereas for $(S, \mathcal{S}) := (S_2, \mathcal{S}_2) = (S_1, \mathcal{S}_1)$ it is called *homogenous* and P is then called a transition (respectively, Markov) kernel on $(\mathcal{S}, S)$.

Consider a family of Markov kernels $(P_\theta : \theta \in \Theta)$ on $(\mathcal{S}_2, S_1)$, where $\Theta \subset \mathbb{R}$ is a compact set, and let $L^1(P_\theta; \Theta) \subset \mathbb{R}^{S_2}$ denote the set of measurable mappings $g : S_2 \to \mathbb{R}$, such that $\int_{S_2} |g(u)| P_\theta(du; s)$ is finite for all $\theta \in \Theta$ and $s \in S_1$.

Definition 3.2 Let $P_\theta \in \mathcal{K}(\mathcal{S}_2, S_1)$, for $\theta \in \Theta$, and let $\mathcal{D} \subset L^1(P_\theta; \Theta)$. We call P_θ differentiable at θ with respect to $\mathcal{D}$, or $\mathcal{D}$*-differentiable* for short, if for any $s \in S_1$ a transition kernel $P'_\theta(\cdot; s) \in \mathcal{M}(S_2, \mathcal{S}_2)$ exists, such that, for any $s \in S_1$ and for all $g \in \mathcal{D}$,

$$\frac{d}{d\theta} \int_{S_2} g(u) P_\theta(du; s) = \int_{S_2} g(u) P'_\theta(du; s). \tag{6}$$

If the left-hand side of (6) equals zero for all $g \in \mathcal{D}$, then we say that P'_θ is not significant.

Recall that $C_b(S)$ denotes the set of continuous bounded mappings from S to $\mathbb{R}$. If $P_\theta \in \mathcal{K}_1(\mathcal{S}_2, S_1)$ is $\mathcal{D}$-differentiable, then $P'_\theta(\cdot; s) \in \mathcal{M}(S_2, \mathcal{S}_2)$ is uniquely defined for any $s \in S_1$, provided that $C_b(S_2) \subset \mathcal{D}$ (see [5]).

If P'_θ exists, then the fact that $P'_\theta(\cdot; s)$ fails to be a probability measure poses the problem of sampling from $P'_\theta(\cdot; s)$. For $s \in S_1$ fixed, we can represent $P'_\theta(\cdot; s)$ by its Hahn-Jordan decomposition as a difference between two probability measures. More precisely, this Hahn-Jordan decomposition is obtained as follows. Let

$$c_{P_\theta}(s) = [P'_\theta]^+(S; s) = [P'_\theta]^-(S; s) \tag{7}$$

and let

$$P^+_\theta(\cdot; s) = [P'_\theta]^+(\cdot; s)/c_{P_\theta}(s), \qquad P^-_\theta(\cdot; s) = [P'_\theta]^-(\cdot; s)/c_{P_\theta}(s);$$

then it holds, for all $g \in \mathcal{D}$, that

$$\int_{S_2} g(u) P'_\theta(du; s) = c_{P_\theta}(s) \left(\int_{S_2} g(u) P^+_\theta(du; s) - \int_{S_2} g(u) P^-_\theta(du; s) \right). \tag{8}$$

For the above line of argument we fixed s. For P^+_θ and P^-_θ to be Markov kernels, we have to consider P^+_θ and P^-_θ as functions in s and have to establish measurability of $P^+_\theta(A; \cdot)$ and $P^-_\theta(A; \cdot)$ for any $A \in \mathcal{S}_2$. This problem is equivalent to showing that $c_{P_\theta}(\cdot)$ in (7) is measurable as a mapping from S_1 to $\mathbb{R}$.

In applications c_{P_θ} is calculated explicitly and its measurability is therefore established case by case. Specifically, in most of the examples presented in this paper,

c_{P_θ} turns out to be a constant and measurability is thus guaranteed. As explained in [10], a general sufficient condition for P'_θ to be a transition kernel is the following: for all $s \in S_1$ it holds that $\sup_{g\in C_b(S_2),|g|\le 1} |\int_{S_2} P'_\theta(du;s)g(u)| < \infty$. In Sect. 4.3 we will show that measurability of c_{P_θ} defined in (7) holds for general state-space S_2 whenever P'_θ is absolutely continuous with respect to another kernel.

To conclude this section, we now introduce the notion of $\mathcal{D}$-derivative, which extends the concept of a weak derivative.

Definition 3.3 Let P_θ be $\mathcal{D}$-differentiable. Any triple $(c_{P_\theta}(\cdot), P^+_\theta, P^-_\theta)$, with $P^\pm_\theta \in \mathcal{K}_1(\mathcal{S}_2, S_1)$ and c_{P_θ} a measurable mapping from S_2 to $\mathbb{R}$, that satisfies (8) is called a *$\mathcal{D}$-derivative* of P_θ. The kernel P^+_θ is called the (normalized) *positive part* of $P_\theta{}'$ and P^-_θ is called the (normalized) *negative part* of $P_\theta{}'$; and $c_{P_\theta}(\cdot)$ is called the *normalizing factor*.

$\mathcal{D}$-derivatives are not unique. To see this, consider $P_\theta \in \mathcal{K}_1(\mathcal{S}_2, S_1)$ with $\mathcal{D}$-derivative $(c_{P_\theta}, P^+_\theta, P^-_\theta)$ and take $Q \in \mathcal{K}(\mathcal{S}_2, S_1)$ so that $\int_{S_2} g(u)Q(du;s)$ is finite for any $g \in \mathcal{D}$ and $s \in S_1$. Set $\tilde{P}^+_\theta = (1/2)P^+_\theta + (1/2)Q$, $\tilde{P}^-_\theta = (1/2)P^-_\theta + (1/2)Q$. Equation (8) implies, for all $g \in \mathcal{D}$ and all $s \in S_1$:

$$\frac{d}{d\theta}\int_{S_2} g(u)P_\theta(du;s) = 2c_{P_\theta}(s)\left(\int_{S_2} g(u)\tilde{P}^+_\theta(du;s) - \int_{S_2} g(u)\tilde{P}^-_\theta(du;s)\right).$$

3.2 Product Rule of Measure-Valued Differentiation

For the finite horizon problem, as stated in (1), the transition kernel P_θ in Definition 3.2 is the n step transition probability of the Markov chain $\{X_\theta(m), m = 0,1,\ldots\}$. In general, it is often very hard to write down the n step transition probability, and studying its differentiability properties is practically impossible. However, the n step transition probability is composed out of one step transition probabilities, that is, transition kernels, which are comparably easier to analyze.

This section establishes the main property of $\mathcal{D}$-differentiable transition kernels, namely, that the product of $\mathcal{D}$-differentiable Markov kernels is again $\mathcal{D}$-differentiable and that the $\mathcal{D}$-derivative can be expressed in terms of the $\mathcal{D}$-derivatives of the transition kernels.

Let P be a Markov kernel on $(\mathcal{S}_2, S_1)$ and Q a Markov kernel on $(\mathcal{S}_1, S_0)$, where $(S_0, \mathcal{S}_0)$ is a measurable Polish space. The product of transition kernels Q, P on $(\mathcal{S}_2, S_0)$ is defined as follows. For $s \in S_0$ and $B \in \mathcal{S}_2$ set $PQ(B;s) = (P \circ Q)(B,s) = \int_{S_1} P(B;z)Q(dz;s)$. If $Q = P$, we set $P^2 = P \circ P$ and $P^n = P^{n-1} \circ P$ for $n \ge 2$. Let $\mathcal{D}_2 \subset L^1(P)$ and $\mathcal{D}_1 \subset L^1(Q)$.

Definition 3.4 Let $\mathcal{D}_2$ be a set of measurable mappings $g: S_2 \to \mathbb{R}$ and let $\mathcal{D}_1 \subset \mathbb{R}^{S_1}$. Transition kernel P_θ is called $(\mathcal{D}_2, \mathcal{D}_1)$-mapping if $\forall g \in \mathcal{D}_2$, $\int_{S_2} g(u)P_\theta(du;\cdot) \in \mathcal{D}_1$. If $\mathcal{D} = \mathcal{D}_1 = \mathcal{D}_2$, then P_θ satisfying the above condition is called $\mathcal{D}$-preserving.

A sufficient condition for $\int g(u)(PQ)(du;s)$ to exist for any $g \in \mathcal{D}_2$ and any $s \in S$, is that P is a $(\mathcal{D}_2, \mathcal{D}_1)$-mapping.

Definition 3.5 Let $P_\theta \in \mathcal{K}(\mathcal{S}_2, S_1)$, $\mathcal{D}^2 \subset L^1(P_\theta;\Theta)$ and $\mathcal{D}^1 \subset \mathbb{R}^{S_1}$ a set of measurable mappings. We call P_θ $(\mathcal{D}^2,\mathcal{D}^1)$-*Lipschitz continuous* if for any $g \in \mathcal{D}^2$ a $K_g \in \mathcal{D}^1$ exists, such that for any $\Delta > 0$ with $\theta + \Delta \in \Theta$ $|\int g(s)P_{\theta+\Delta}(ds;\cdot) - \int g(s)P_\theta(ds;\cdot)| \le \Delta K_g$. If $\mathcal{D} = \mathcal{D}^2 = \mathcal{D}^1$, then we call P_θ simply $\mathcal{D}$-*Lipschitz continuous*.

The following theorem presents the key technical result of this section.

Theorem 3.1 *Let $((S_i,\mathcal{S}_i) : 0 \le i \le n)$ be a sequence of Polish measurable spaces. For $1 \le i \le n$, let $P_{\theta,i}$ be a transition kernel on $(\mathcal{S}_i, S_{i-1})$, such that $P_{\theta,i}$ is $\mathcal{D}_i$-differentiable. Furthermore, set $\mathcal{D}_0 = \mathbb{R}^{S_0}$.*

We introduce the following assumptions for each i, $1 \le i \le n$:

(A0) *if $g, f \in \mathcal{D}_i$ then it holds that $f + g \in \mathcal{D}_i$,*
(A1) *$P_{\theta,i}$ is a $(\mathcal{D}_i,\mathcal{D}_{i-1})$-mapping,*
(A2) *$P_{\theta,i}$ is $(\mathcal{D}_i,\mathcal{D}_{i-1})$-Lipschitz continuous,*
(A3) *$P_{\theta,i}$ is $\mathcal{D}_i$-differentiable such that $P'_{\theta,i} \in \mathcal{K}(S_i, S_{i-1})$ and $P'_{\theta,i}$ is a $(\mathcal{D}_i,\mathcal{D}_{i-1})$-mapping.*

The following statements hold true:

(i) *Under Assumptions* (A0), (A1), (A2), $\prod_{i=1}^n P_{\theta,i}$ *is $(\mathcal{D}_n,\mathcal{D}_0)$-Lipschitz continuous.*
(ii) *Under Assumptions* (A0), (A1), (A2), (A3) *the following product rule holds:*

$$\left(\prod_{i=1}^n P_{\theta,i}\right)' = \sum_{j=1}^n \prod_{i=j+1}^n P_{\theta,i} P'_{\theta,j} \prod_{i=1}^{j-1} P_{\theta,i}.$$

Following the line of proof of Theorem 3.1, one obtains the following chain rule of differentiation.

Corollary 3.1 *Consider a $\mathcal{D}$-differentiable Markov kernel P_θ such that $P'_\theta \in \mathcal{K}$. Let $g_\theta \in \mathcal{D}$ and assume that $K_g \in \mathcal{D}$ exists, such that for any $\Delta \in \mathbb{R}$ with $\theta + \Delta \in \Theta$ $|g_{\theta+\Delta}(s) - g_\theta(s)| \le \Delta K_g(s)$. If g_θ is differentiable at θ, then*

$$\frac{d}{d\theta}\int g_\theta(u)P_\theta(du;s) = \int \left(\frac{d}{d\theta} g_\theta(u)\right) P_\theta(du;s) + \int g_\theta(u)P'_\theta(du;s), \tag{9}$$

for any $s \in S$.

Remark 3.1 When the performance function depends explicitly on θ the first term in (9) is recognizably the so-called *IPA term*, and the corresponding integrability assumption is given as a weak $\mathcal{D}$ Lipschitz continuity assumption. In particular, when the kernel is independent of θ the corollary recovers the usual IPA formulation. It is worthwhile to notice that the path-wise analysis common to SPA/IPA formulations requires explicit construction of the trajectories to evaluate the propagation of the perturbations: our formalism implicitly deals with this propagation through a simple chain rule of differentiation.

Often, one is interested in evaluating expected values of an entire trajectory rather than a particular n step transition, see (1). As detailed in [5], the product rule of measure-valued differentiation can be applied to this problem too.

Theorem 3.1 establishes a product rule for measure-valued differentiation of Markov kernels. In order to verify whether a given Markov kernel satisfies the conditions in Theorem 3.1, it is often helpful to separate the parts of the transition kernel that depend on θ and those that are independent of θ. We illustrate this conditioning approach with the following example.

Example 3.1 For $\Theta = [0, 1]$, let $\eta_\theta \in \{0, 1\}$ be Bernoulli-θ-distributed on $S_\eta = \{0, 1\}$, with $\mathbb{P}(\eta_\theta = 0) =: \mu_\theta(0) = \theta$, $\mathbb{P}(\eta_\theta = 1) =: \mu_\theta(1) = 1 - \theta$. For any $g = (g_0, g_1) \in \mathbb{R}^2$ it holds that

$$\frac{d}{d\theta} \int_{S_\eta} g_s \mu_\theta(ds) = g_0 - g_1 = \int_{S_\eta} g_s \delta_0(ds) - \int_{S_\eta} g_s \delta_1(ds),$$

where δ_y denotes the Dirac measure on y. Thus, μ_θ has $\mathbb{R}^2$-derivative $(1, \delta_0, \delta_1)$.

Let $\{X_\theta(n)\}$ denote the queue-length processes of a Markovian queueing network. Denote the transition kernel of $\{X_\theta(n)\}$ by P_θ, where θ is a routing parameter. The routing decision is made as follows. If, at the nth state transition, a customer leaves a particular sever of the network, a Bernoulli-(θ)-distributed random variable $\eta_\theta(n)$ is generated independent of everything else. For $\eta_\theta(n) = 0$ the customer is routed to a particular server, say j, and for $\eta_\theta(n) = 1$ he/she is routed to a server, say j', with $j \neq j'$. Using the fact that $\{\eta_\theta(n)\}$ is an i.i.d. sequence, we can draw a sample of $\eta_\theta(n)$ at each transition. Let $Q(\cdot; s, \cdot)$ denote the transition kernel of $X_\theta(n)$ given that $\eta_\theta(n) = s$ and let $\mathcal{D}$ be the set of all g such that, for any possible queue-length vector x and $s = 0, 1$,

$$\mathbb{E}\big[|g(X_\theta(n+1))| |X_\theta(n) = x, \eta_\theta = s\big] < \infty. \tag{10}$$

Then, P_θ is $\mathcal{D}$-differentiable. More specifically, for any $g \in \mathcal{D}$, it holds that

$$\begin{aligned}\frac{d}{d\theta} \int g(u) P_\theta(du; s) &= \frac{d}{d\theta} \int_{S_\eta} \int g(u) Q(du; \eta, s) \mu_\theta(d\eta)\\ &= \int g(u) Q(du; 0, s) - \int g(u) Q(du; 1, s)\\ &= \int g(u) P_0(du; s) - \int g(u) P_1(du; s).\end{aligned}$$

Hence, a $\mathcal{D}$-derivative of P_θ can be obtained from $(1, P_0, P_1)$. Moreover, the $\mathcal{D}$-derivative is independent of θ and P_θ is thus $\mathcal{D}$-Lipschitz continuous (for a proof use the mean-value theorem). Linearity of the expected value in (10) implies that for $f, g \in \mathcal{D}$ it holds that $f + g \in \mathcal{D}$. Hence, provided that $\int g(u) P_\theta(du; \cdot) \in \mathcal{D}$ for any $g \in \mathcal{D}$ and $\theta \in [0, 1]$, the product rule applies to P_θ.

To emphasize the potential benefits of the MVD approach, we stress that from the simple formulas for the weak derivative of a Bernoulli random variable it is now possible to reconstruct the MVD formulas for the routing sensitivities in the whole network, via the product rule.

The above conditioning approach can be interpreted as a particular kind of conditioning within the SPA setting (see Sect. 2.1), although MVD does not yield a pathwise estimator, but a closed formula for the distributions. In Sect. 5 we specifically deal with the construction of various estimators from MVD formulas.

4 Setting the Product Rule to Work

In this section, we discuss various meaningful ways of interpreting the conditions in Theorem 3.1. Simple examples will be given to illustrate the situations we have in mind. For the sake of simplicity, consider homogeneous Markov chains and denote the state-space by $(S, \mathcal{S})$. To simplify the notation, drop the explicit dependence on the state-space whenever this causes no confusion. For example, we will write C_b instead $C_b(S)$ for the set of bounded continuous functions.

4.1 Bounded Performance Functions

As a first choice for $\mathcal{D}$ take $\mathcal{D}^0$: the set of bounded measurable mappings, which satisfies (A0).

Lemma 4.1 *Let P_θ be a Markov kernel that is $\mathcal{D}^0$-differentiable on Θ with $\mathcal{D}^0$-derivative $((c_{P_\theta}(s), P_\theta^+(\cdot; s), P_\theta^-(\cdot; s)) : s \in S)$. If $\sup_{\theta\in\Theta} c_{P_\theta}(\cdot) \in \mathcal{D}^0$, then $\left(P_\theta^n\right)' = \sum_{j=1}^n P_\theta^{n-j} P_\theta' P_\theta^{j-1}$.*

Example 4.1 Let $X_\theta(n)$ be the discrete-time queue length process of an M/M/1/m queue with m buffer places, arrival rate λ and service rate θ, with $\theta \ge a > 0$. The transition kernel is then given in matrix form by

$$P_\theta = \begin{bmatrix} 0 & 1 & 0 & & & \\ \frac{\lambda}{\lambda+\theta} & 0 & \frac{\theta}{\lambda+\theta} & 0 & & \\ 0 & \frac{\lambda}{\lambda+\theta} & 0 & \frac{\theta}{\lambda+\theta} & 0 & \cdots \\ & & & \ddots & & \\ & & & & 1 & 0 \end{bmatrix}.$$

Let $c_{P_\theta}(k) = (\lambda + \theta)^{-2}\mathbf{1}_{\{0<k\le m\}}$, then a weak derivative of $P_\theta(\cdot; k)$ can be obtained as follows. For $k = 0$ and $k = m + 1$, $(c_{P_\theta}(k), P_\theta^+(\cdot; k), P_\theta^-(\cdot; k)) = (0, P_\theta(\cdot; k), P_\theta(\cdot; k))$ and for $0 < k \le m$ $(c_{P_\theta}(k), P_\theta^+(\cdot; k), P_\theta^-(\cdot; k)) = ((\lambda/(\lambda + \theta)^2), \delta_{k+1}(\cdot), \delta_{k-1}(\cdot))$. Since $\theta \ge a > 0$, $\sup_{\theta\in\Theta} c_{P_\theta}(\cdot) \in C_b$ and Lemma 4.1 yields, for example, a closed form expression for the derivative of any moment of the queue length at the nth state with respect to the service rate.

For this setup the $\mathcal{D}$-derivative can be represented in a concise form through matrix notation. To see this, define the matrix C_{P_θ} and the matrices P^+, P^- by

$$C_{P_\theta} = \begin{bmatrix} 0 & & & \\ & \frac{1}{(\lambda+\theta)^2} & & \\ & & \ddots & \\ & & & \frac{1}{(\lambda+\theta)^2} & \\ & & & & 0 \end{bmatrix}, \qquad P^+ = \begin{bmatrix} 0 & 0 & & & \\ & 0 & 1 & & \\ & & \ddots & & \\ & & & 0 & 1 \\ & & & 0 & 0 \end{bmatrix},$$

$$P^- = \begin{bmatrix} 0 & 0 & & & \\ & 1 & 0 & & \\ & & \ddots & & \\ & & & 1 & 0 \\ & & & 0 & 0 \end{bmatrix};$$

then,

$$\frac{d}{d\theta} P_\theta = C_{P_\theta}(P^+ - P^-),$$

and (with slight abuses of notation) the triple (C_{P_θ}, P^+, P^-) may serve as matrix-valued $\mathcal{D}$-derivative of P_θ. The statement of Lemma 4.1 then reads

$$\frac{d}{d\theta} P_\theta^n = \sum_{j=1}^{n} P_\theta^{n-j} C_{P_\theta} P^+ P_\theta^{j-1} - \sum_{j=1}^{n} P_\theta^{n-j} C_{P_\theta} P^- P_\theta^{j-1}.$$

4.2 Performance Functions Bounded by a Polynomial

It is often too restrictive in applications to assume that the sample performance is bounded ($g \in \mathcal{D}^0$). A convenient set of functions is the set $\mathcal{D}^p$ of polynomially bounded performance functions given by

$$\mathcal{D}^p = \left\{ g : S \to \mathbb{R} \,\middle|\, |g(x)| \le \sum_{i=0}^{p} \kappa_i \|x\|^i, \kappa_i \in \mathbb{R}, 0 \le i \le p \right\}, \tag{11}$$

for some $p \in \mathbb{N}$, where $\|\cdot\|$ denotes a norm on S (assuming that S is indeed equipped with a norm). Most cases of interest in applications fall within this setting. Note that $\mathcal{D}^p$ satisfies (A0) and that $\mathcal{D}^p \subset L^1(P_\theta : \Theta)$ if $P_\theta(\cdot; s)$ has finite pth moment for any $s \in S$ and $\theta \in \Theta$. The above definition recovers $\mathcal{D}^0$ as the set of bounded functions.

Lemma 4.2 *Let $p \in \mathbb{N}$. Consider a (homogeneous) Markov kernel P_θ with finite pth moment for any $s \in S$ and $\theta \in \Theta$. Assume that P_θ is $\mathcal{D}^p$-differentiable on Θ with $\mathcal{D}^p$-derivative $((c_{P_\theta}(s), P_\theta^+(\cdot; s), P_\theta^-(\cdot; s)) : s \in S)$. If P_θ' is $\mathcal{D}^p$-preserving and a $K(\cdot) \in \mathcal{D}^p$ exists, such that* $\sup_{\theta\in\Theta}(c_{P_\theta}(\cdot) \int (1 + \|s\|^p) P_\theta^{\pm}(ds; \cdot)) \le K(\cdot)$, *then* $(P_\theta^n)' = \sum_{j=1}^{n} P_\theta^{n-j} P_\theta' P_\theta^{j-1}$.

Example 4.2 Let $X_\theta(n)$ denote the n-th waiting time at a GI/F_θ/1 queue with generally distributed inter-arrival times. Let η_θ denote the exponential distribution with mean $1/\theta$ and $\Gamma(2,\theta)$ the gamma$(2,\theta)$ distribution. The service times are governed by the distribution $F_\theta = \theta\eta_{\theta_0} + (1-\theta)\Gamma(2,\theta_0)$, $\theta \in \Theta = [0,1]$, that is, with probability θ the service time is exponentially distributed with mean θ_0 and with probability $1-\theta$ it is distributed like the sum of two independent exponentials with mean θ_0 each. We have $S = \mathbb{R}$ and we take the usual norm on $\mathbb{R}$ for $\|\cdot\|_S$. Observe that F_θ is $\mathcal{D}^p$-differentiable for any p and a $\mathcal{D}^p$-derivative is given by

$$(1, \eta_{\theta_0}, \Gamma(2,\theta_0)), \tag{12}$$

which is independent of θ. Let $\{A(n)\}$ be the i.i.d. sequence of inter-arrival times and $\{S_\theta(n)\}$ the i.i.d. sequence of service times, respectively. Lindley's recursion yields $X_\theta(n+1) = \max(X_\theta(n) + S_\theta(n) - A(n+1), 0)$, $n > 1$, and $X_\theta(1) = 0$. As performance function, take the pth moment of the waiting times (which is not in $\mathcal{D}^0$). Let $G(\cdot)$ denote the distribution of the inter-arrival time and assume that the first p moments of G are finite. For $w > 0$, the transition kernel for the waiting times is given by

$$P_\theta((0,w];v) = \int_0^\infty \int_{s+v-w}^{s+v} G(da) F_\theta(ds) =: \int Q((0,w];s,v) F_\theta(ds),$$

$$P_\theta(\{0\};v) = \int_0^\infty \int_{s+v}^{\infty} G(da) F_\theta(ds) =: \int Q(\{0\};s,v) F_\theta(ds).$$

For any $g \in \mathcal{D}^p$, then it holds that $\int g(u) P_\theta(du;v) \in \mathcal{D}^p$ and P_θ is thus $\mathcal{D}^p$-preserving.

The first step is to calculate the $\mathcal{D}^p$-derivative of P_θ. For any $v \geq 0$ and $g \in \mathcal{D}^p$, $\int g(s) Q(dr;\cdot,v)$ is again in $\mathcal{D}^p$ and since F_θ is $\mathcal{D}^p$-differentiable it easily follows that $P_\theta(\cdot;v)$ is $\mathcal{D}^p$-differentiable. A $\mathcal{D}^p$-derivative of F_θ is given in (12) and a $\mathcal{D}^p$-derivative of P_θ can therefore be obtained from

$$P^+((0,w];v) = \begin{cases} \int_0^\infty Q((0,w];s,v)\eta_{\theta_0}(ds), & w > 0, \\ \int_0^\infty Q(\{0\};s,v)\eta_{\theta_0}(ds), & w = 0, \end{cases}$$

$$P^-((0,w];v) = \begin{cases} \int_0^\infty Q((0,w];s,v)\Gamma(2,\theta_0)(ds), & w > 0, \\ \int_0^\infty Q(\{0\};s,v)\Gamma(2,\theta_0)(ds), & w = 0, \end{cases}$$

with $c_{P_\theta} = 1$. Note that this simple calculation implies that $P'_\theta = P^+ - P^-$ is a transition kernel.

Longer service times lead to longer waiting times, which implies the following chain of inequalities, for any $\theta \in [0,1]$, $\int (1+u)^p P^+(du;v) \leq \int (1+u)^p P_\theta(du;v) \leq \int (1+u)^p P^-(du;v) = \int (1+u)^p P_1(du;v)$, for $v \geq 0$. Note that $\int (1+u)^p P_1(du;\cdot) =: K(\cdot) \in \mathcal{D}^p$. Hence, P'_θ is $\mathcal{D}^p$-preserving. Moreover, elaborating on the fact that $c_{P_\theta} = 1$ and that $P^\pm$ are independent of θ, it readily follows that

$$\sup_{\theta\in\Theta} \int (1+u)^p P^\pm(du;v) = \int (1+u)^p P^\pm(du;v) \leq K(v).$$

Hence, Lemma 4.2 yields, for example, a closed form expression for the derivative of the pth moment of the nth waiting time at a GI/F_θ/1 queue.

4.3 Markov Kernels with Differentiable Densities

As already illustrated in Sect. 2.3, the analysis of derivatives of stochastic systems simplifies when the distributions involved have densities that are differentiable as functions of θ. In this section, we will illustrate how the conditions for the product rule of measure-valued differentiation simplify under the presence of differentiable densities.

For $P, Q \in \mathcal{K}_1$, let P be absolutely continuous with respect to Q, in symbols: $P \ll Q$. This implies that the Radon-Nikodym derivative of $P(\cdot; s)$ with respect to $Q(\cdot; s)$. exists for all s, and we denote it by $[dP/dQ](r; s)$ with $r, s \in S$. If P_θ is absolutely continuous with respect to Q, then the positive and negative part of the $\mathcal{D}$-derivative of P_θ is given through integrating the positive and negative parts of the derivative of $[dP'_\theta/dQ](r; s)$, and the corresponding the normalizing factor is measurable. The precise statement is given in the following lemma.

Lemma 4.3 *Let $P_\theta, Q \in \mathcal{K}_1$, for $\theta \in \Theta$. Assume that P_θ is $\mathcal{D}$-differentiable at θ and $P'_\theta \ll Q$. Then, $P'_\theta \in \mathcal{K}$ and $(c_{P_\theta}, P^+_\theta, P^-_\theta)$ is a $\mathcal{D}$-derivative of P_θ, with $c_{P_\theta}(s) = \int_S \max(0, [\frac{dP'_\theta}{dQ}](r; s)) Q(dr; s)$, $s \in S$, for any $A \in \mathcal{S}$ and $s \in S$,*

$$P^+_\theta(A; s) = \frac{1}{c_{P_\theta}(s)} \int_A \max\left(0, \left[\frac{dP'_\theta}{dQ}\right](r; s)\right) Q(dr; s),$$

$$P^-_\theta(A; s) = \frac{1}{c_{P_\theta}(s)} \int_A \max\left(0, -\left[\frac{dP'_\theta}{dQ}\right](r; s)\right) Q(dr; s).$$

The following lemma establishes sufficient conditions for the product rule to hold in the presence of domination.

Lemma 4.4 *Let $p \in \mathbb{N}$. Consider a (homogeneous) Markov kernel P_θ, with finite pth moment for any $s \in S$ and $\theta \in \Theta$. Assume that P_θ is $\mathcal{D}^p$-preserving and $\mathcal{D}^p$-differentiable on Θ, and that $P'_\theta \ll P_\theta$. If P'_θ is $\mathcal{D}^p$-preserving and a $K(\cdot) \in \mathcal{D}^p$ exists such that*

$$\sup_{\theta \in \Theta} \int (1 + \|s\|^p) \left| \left[\frac{dP'_\theta}{dP_\theta}\right](ds; \cdot) \right| P_\theta(ds; \cdot) \le K(\cdot),$$

then

$$(P^n_\theta)' = \sum_{j=1}^{n} P^{n-j}_\theta P'_\theta P^{j-1}_\theta.$$

The key to applying Lemma 4.4 is to compute the Radon-Nikodym derivative of P'_θ with respect to P_θ. In applications, P'_θ is typically of rather complex structure and computing the Radon-Nikodym derivative of P'_θ with respect to P_θ leads to cumbersome calculations, as will be illustrated in Example 4.3. However, using a conditioning argument, the assumption in Lemma 4.4 can be restated in terms of conditions that are easier to verify.

Example 4.3 Let $X_\theta(n)$ denote the n-th waiting time at a GI/M/1 queue. Let $\{A(n)\}$ be the sequence of interarrival times and $\{S_\theta(n)\}$ the sequence of exponentially distributed service times with mean $1/\theta$, respectively. Let $\Theta = [a, b] \subset (0, \infty)$. Denote the distribution of $S_\theta(n)$ by η_θ and the corresponding Lebesgue density by $f_\theta^S(x) = \theta e^{-\theta x}$. Let $A(n)$ have a finite pth moment and let f^A denote the Lebesgue density of the interarrival times. As performance measure of interest, consider the pth moment of the waiting time. Let P_θ denote the transition kernel of $\{X_\theta(n)\}$. Following the line of thought in Example 4.2, for any $w > 0$, $v \geq 0$, the transition kernel for the waiting times is given by

$$P_\theta((0,w];v) = \int_0^\infty \int_{v+x-w}^{v+x} f^A(a) f_\theta^S(x) da dx =: \int_0^\infty R((0,w];x,v) f_\theta^S(x) dx,$$

$$P_\theta(\{0\};v) = \int_0^\infty \int_{v+x}^{\infty} f^A(a) f_\theta^S(x) da dx =: \int_0^\infty R(\{0\};x,v) f_\theta^S(x) dx.$$

The exponential distribution is $\mathcal{D}^p$-differentiable for any p (see Example 2.1). Differentiating P_θ with respect to θ yields

$$P'_\theta((0,w];v) = \int_0^\infty R((0,w];x,v)(1-\theta x) f_\theta^S(x) dx, \quad w > 0,$$

$$P'_\theta(\{0\};v) = \int_0^\infty R(\{0\};x,v)(1-\theta x) f_\theta^S(x) dx.$$

A $\mathcal{D}^p$-derivative of P_θ can be obtained from setting $c_{P_\theta} = 1/\theta$ and, for any measurable set A,

$$P_\theta^+(A;v) = \int_0^\infty R(A;x,v) f_\theta^S(x) dx = P_\theta(A;v).$$

For $w > 0$,

$$P_\theta^-((0,w];v) = \int_0^\infty R((0,w];x,v) h_\theta^S(x) dx = \int_0^\infty \int_{v+x}^{v+x-w} f^A(a) h_\theta^S(x) da dx,$$

$$P_\theta^-(\{0\};v) = \int_0^\infty R(\{0\};x,v) h_\theta^S(x) ds = \int_0^\infty \int_{v+x}^{\infty} f^A(a) h_\theta^S(x) da dx,$$

where h_θ denotes the density of the gamma(2,θ) distribution. Let $\Theta = [a, b]$, with $a > 0$.

We now show that the product rule of measure-valued differentiation applies to P_θ. For any $g \in \mathcal{D}^p$, it holds that

$$H_g(x,v) = \int_0^\infty g(u) R(du;x,v) = \int_{v+x}^\infty g(v+x-a) f^A(a) da,$$

assuming for the sake of simplicity that $g(0)=0$. Because the inter-arrival times have finite pth moment (by assumption) it is easily verified that for $g\in\mathcal{D}^p$ the mapping $\int H_g(x,\cdot)f_\theta^S(x)dx$ is in $\mathcal{D}^p$ and P_θ is hence $\mathcal{D}$-preserving. Following the same line of argument, it is easily seen that $\int H_g(x,\cdot)h_\theta^S(x)dx$ is in $\mathcal{D}^p$ for any $g\in\mathcal{D}^p$ and thus $P'_\theta=(1/\theta)(P_\theta^+-P_\theta^-)$ is $\mathcal{D}^p$-preserving.

Note that $\eta'_\theta\ll\eta_\theta$ for any $\theta\in\Theta=[a,b]$. Then, replacing η_θ by the corresponding density f_θ^S,

$$\sup_{\theta\in[a,b]}\int_0^\infty(1+u^p)\left|\frac{\frac{d}{d\theta}f_\theta^S(u)}{f_\theta^S(u)}\right|f_\theta^S(u)du\le\frac{1}{a}\int_0^\infty(1+u^p)(1+bu)f_a^S(u)du,$$

which is finite for any $p\in\mathbb{N}$. In accordance with Lemma 4.2, for any $g\in\mathcal{D}^p$, with $p\in\mathbb{N}$, the product rule applies to P_θ and yields a closed-form expression for the derivative of the pth moment of the waiting time.

5 Gradient Estimation

While MVD offers a methodology that helps to establish a closed formula for (1), in practice one wishes to construct an estimator based on observations (or simulations) of the underlying Markov process. Let $(P_{\theta,i}:1\le i\le n)$ be a family of $\mathcal{D}$-differentiable Markov kernels on $(S,\mathcal{S})$. The product rule of measure-valued differentiability yields

$$\begin{aligned}
&\frac{d}{d\theta}\mathbb{E}[g(X_\theta(n),\dots,X_\theta(1))]\\
&\quad=\frac{d}{d\theta}\int g(s_n,\dots,s_1)\prod_{i=1}^{n}P_{\theta,i}(ds_i;s_{i-1})\\
&\quad=\sum_{j=1}^{n}\iiint g(s_n,\dots,s_1)\prod_{i=j+1}^{n}P_{\theta,i}(ds_i;s_{i-1})P'_{\theta,j}(ds_j;s_{j-1})\prod_{i=1}^{j-1}P_{\theta,i}(ds_i;s_{i-1}),
\end{aligned}$$

with $s_0\in S$. How to transform the above into an unbiased estimator?

5.1 Phantom Estimators

In this section, we establish sufficient conditions for phantom type estimators to be unbiased. Let the conditions in Theorem 3.1 be in force. Then one can express $P'_\theta(ds_j;s_{j-1})$ in terms of the normalized difference between two expectations, or $(c_{P_\theta}(s_{j-1}),P^+_{\theta,j}(ds_j;s_{j-1}),P^-_{\theta,j}(ds_j;s_{j-1}))$. The product form can be rewritten in terms of the processes as follows:

$$\begin{aligned}
\frac{d}{d\theta}\mathbb{E}[g(X_\theta(n),\dots,X_\theta(1))]=\sum_{j=1}^{n}\Big(&\mathbb{E}\big[c_{P_\theta}(X_\theta(j-1))g(X^+_{\theta,j}(n),\dots,X^+_{\theta,j}(1))\big]\\
&-\mathbb{E}\big[c_{P_\theta}(X_\theta(j-1))g(X^-_{\theta,j}(n),\dots,X^-_{\theta,j}(1))\big]\Big),
\end{aligned}\tag{13}$$

where the processes $\{X^{\pm}_{\theta,j}(i)\}$ are Markov chains that follow the transition kernels $P_{\theta,i}(ds_i; s_{i-1}), i \neq j$, and where $\mathbb{P}(X^{\pm}_{\theta,j}(j) \in \cdot | X^{\pm}_{\theta,j}(j-1) = s_{j-1}) = P^{\pm}_{\theta,j}(\cdot; s_{j-1})$. The chains $\{X^{\pm}_{\theta,j}(i)\}$ are called *phantoms* in the literature.

Equation (13) has the following interpretation: the processes $\{X^{\pm}_{\theta,j}(n), j = 1, \dots, N\}$ follow the transition kernel of the process $\{X_\theta(n)\}$ up to $n = j-1$. Next, $\mathbb{P}(X^{\pm}_{\theta,j}(j) \in \cdot | X^{\pm}_{\theta,j}(j-1) = x)$ follows the kernel $P^{\pm}_\theta(\cdot; x)$. After this transition, again the one step transition kernel of the processes $X^{\pm}_{\theta,j}$ follow $P_\theta(\cdot; x)$.

Example 5.1 Consider a standard periodic review inventory model with backlog. Consecutive demands $\{D(n)\}$ are assumed continuous with Lebesgue density $f(\cdot)$, so that the inventory level $\{X_\theta(n)\}$ at the review epochs is Markovian, $X_\theta(0) = 0$ and for $n \geq 0$:

$$X_\theta(n+1) = \begin{cases} X_\theta(n) - D(n+1), & \text{if } X_\theta(n) - D(n+1) \geq \theta, \\ S, & \text{otherwise,} \end{cases}$$

with $D(n)$ an i.i.d. sequence, and where the control variable θ represents the threshold for the ordering policy and S is the total storage capacity. Call P_θ the corresponding kernel. The cost per period is: $\tilde{g}(X_\theta(n), D(n)) = h(X_\theta(n) - D(n))\mathbf{1}_{\{D(n) < X_\theta(n)\}} + p(D(n) - X_\theta(n))\mathbf{1}_{\{D(n) > X_\theta(n)\}} + K\mathbf{1}_{\{D(n) > X_\theta(n) - \theta\}}$, where h is unit holding cost, K is ordering cost and p is a backlog penalty. Define the integrated cost per period at state $X_\theta(n) = x$ by $g(x) = \mathbb{E}[\tilde{g}(X_\theta(n), D(n)) | X_\theta(n) = x]$. The finite horizon cost is

$$J(\theta) = \sum_{n=1}^{N} \mathbb{E}\big[g(X_\theta(n))\big],$$

This cost function is not a.s. Lipschitz continuous. Moreover, because θ is a threshold parameter then actually $\frac{d}{d\theta} g(X_\theta(n)) = 0$, a.s., so that

$$\mathbb{E}\left[\sum_{n=1}^{N} \frac{d}{d\theta} g(X_\theta(n))\right] \neq \frac{d}{d\theta} \sum_{n=1}^{N} \mathbb{E}\big[g(X_\theta(n))\big].$$

The problem has been studied using the SPA path-wise methodology, see [2]. It is clear that, for any bounded and continuous function $g \in \mathcal{D}^0$ and all $x \in (\theta, S]$,

$$\mathbb{E}[g(X_\theta(n+1)) | X_\theta(n) = x] = \int_0^{x-\theta} g(x-y) f(y) dy + g(S)(1 - F(x-\theta)),$$

where f is the Lebesgue density and F the cumulative distribution function of the demand $D(n)$. The derivative is calculated directly, yielding: $d\mathbb{E}[g(X_\theta(n+1)) | X_\theta(n)] / d\theta = f(X_\theta(n) - \theta)(g(S) - g(\theta)) = c_\theta(X_\theta(n))\mathbb{E}[g(X^+_\theta(n+1)) - g(X^-_\theta(n+1))]$, where $c_\theta(x) = f(x - \theta)$ for $x \in \mathbb{R}$ and the random variables $X^{\pm}_\theta(n+1)$ are concentrated at the mass points S and θ respectively (note that $c_\theta(\cdot)$ is measurable). Because $\sup_{\theta \in [0,S]} c_\theta(\cdot) = \sup_{\theta \in [0,S]} f((\cdot) - \theta) \in \mathcal{D}^0$, the product rule of measure-valued differentiation applies to P_θ, see Lemma 4.1. Specifically, the product rule for $\mathcal{D}$-differentiability prescribes defining the processes $\{X^{\pm}_{\theta,j}(i)\}$ as follows: the first

$j-1$ transitions are governed by the kernel P_θ as the inventory process itself. Next $X^+_{\theta,j}(j)=S$, $X^-_{\theta,j}(j)=\theta$ and the rest of the transitions are again governed by P_θ. The product rule of $\mathcal{D}$-differentiation, see Theorem 3.1, yields

$$\begin{aligned}\frac{d}{d\theta}J(\theta) &= \sum_{j=1}^{N}\int\cdots\int\sum_{n=1}^{N}g(x_n)\prod_{i=j+1}^{N}P_\theta(dx_i;x_{i-1})P'_\theta(dx_j;x_{j-1})\\ &\quad\times\prod_{i=1}^{j-1}P_\theta(dx_i;x_{i-1})\\ &= \mathbb{E}\left[\sum_{j=1}^{N}c_\theta(X_\theta(j-1))\sum_{n=j}^{N}\big(g(X^+_{\theta,j}(n))-g(X^-_{\theta,j}(n))\big)\right].\end{aligned}$$

The SPA estimator of [2] is an instance of the above processes, using common random variables for the past history up to transition j. In this example, decoupling occurs because $f(X_\theta(j)-\theta)$ is independent of $X^{\pm}_{\theta,j}(i), i>j$, which allows for several implementations.

Sections 4.1 and 4.2 studied different choices for the space functions $\mathcal{D}$. Obviously, the minimal condition on $\mathcal{D}$ is absolute integrability with respect to $P^n\mu$ for any n, where μ is the initial distribution of the Markov chain. Consider a Markov chain $\{X_\theta(n)\}$ with transition kernel P_θ, and use the notation P^n_θ to indicate the n-step transition probability. In what follows, assume that the initial distribution μ of the Markov chain is fixed. Let

$$\mathcal{D}_\mu=\left\{g:S\to\mathbb{R}|\forall n\in\mathbb{N}\forall\theta\in\Theta:\int|g(u)|(P^n_\theta\mu)(du)<\infty\right\},$$

or in terms of random variables,

$$\mathcal{D}_\mu=\left\{g:S\to\mathbb{R}|\forall n\in\mathbb{N}\forall\theta\in\Theta:\int\mathbb{E}_s[|g(X_\theta(n))|]\mu(ds)<\infty\right\},$$

where $\mathbb{E}_s$ denotes the expected value conditioned on the event $X(0)=s$, for $s\in S$. In the presence of domination it is possible to state a sufficient condition for the product rule to hold on $\mathcal{D}_\mu$.

Lemma 5.1 *For $\theta\in\Theta$, let $\{X_\theta(n)\}$ be a homogeneous Markov chain with transition kernel P_θ, where Θ is a neighborhood of θ^* and assume that $P_\theta\ll P_{\theta^*}$. Let P_θ be $\mathcal{D}_\mu$-differentiable, such that $P^+_\theta, P^-_\theta\ll P_{\theta^*}$. Assume that*

(a) *For any n and any $s\in S$, $\mathbb{E}_s[|g(X_\theta(n))|]<\infty$.*

(b) *For any n, $\mathbb{E}_s[\sup_{\theta\in\Theta}|g(X_\theta(n))||\frac{dP'_\theta}{dP_\theta}(X_\theta(n),X_\theta(n-1))|]<\infty$.*

Then,

$$\frac{d}{d\theta}\mathbb{E}_s[g(X_\theta(n))]=\sum_{j=1}^{n}\mathbb{E}_s[c_{P_\theta}(X_\theta(j-1))(g(X^+_{\theta,j}(n))-g(X^-_{\theta,j}(n)))].\tag{14}$$

5.2 Single-Run Estimation

In this section we consider the question of single-run estimation versus estimators that require parallel simulations. Observe that in practical engineering situations only a single sample path of a system may be available. In such a situation an estimator requiring parallel simulation may not be feasible (or only available via cut-and-past techniques at high computational costs). In addition, when one is willing to carry out parallel simulations, one may as well carry out parallel simulations in order to estimate the derivative via finite differences, thus resorting to a biased estimator for the gradient. A thorough discussion of these issues is beyond the scope of this paper and we focus on unbiased gradient estimators in the case when single-run estimators and estimators requiring parallel simulations are applicable.

Because the efficiency of an estimator measures the trade off between speed and variance, it is sometimes assumed that a single-run estimator, which only needs observations of the sample path of the process, is *a priori* preferable to other estimators. Both SPA and SF are single-run estimators. In our first example we examine the variance of several estimators, where the single-run versions may actually behave worse than a parallel experiment.

Example 5.2 Consider the following simple gradient-estimation problem. Let X_θ follow an exponential distribution with mean $1/\theta$. Suppose we are interested in estimating $d\mathbb{E}[g(X_\theta)]/d\theta$, where g is of the form $g(x) = x^p$ for some $p \in \mathbb{N}$. In this case no discontinuities occur and the IPA estimator applies. In order to obtain the IPA estimator, note that $1/\theta$ is a scaling parameter of the exponential distribution, which implies $dX_\theta/d\theta = -X_\theta/\theta$. Hence, taking the sample path derivative of $g(X_\theta)$ with respect to θ one arrives at the estimator $\mathbb{E}[-(p/g)(X_\theta)]$, where we have used the fact that $g(x) = x^p$ for some $p \in \mathbb{N}$.

The SF estimator follows from differentiating the Lebesgue density of X_θ and reads $\mathbb{E}[g(X_\theta)(1 - \theta X_\theta)]$.

Using the results of Example 2.1, the MVD derivative can be estimated by simulating two random variables in parallel processes (called "phantoms"). A phantom estimator can be obtained from $(1/\theta)\mathbb{E}[g(X_\theta) - g(Y_\theta)]$, where Y_θ follows a gamma $(2, \theta)$ distribution. This expression offers the possibility for variance reduction via common random numbers. A natural choice is to consider $(1/\theta)\mathbb{E}[g(X_\theta(1)) - g(X_\theta(1) + X_\theta(2))]$ as estimator, where $X_\theta(1)$ and $X_\theta(2)$ are independent and exponentially distributed random variables with mean $1/\theta$. This is referred to as the "Coupled Phantom" estimator.

All estimators are unbiased and Table 1 gives the variance per estimator for different choices of p in the performance function. The Phantom estimator is obtained using independent random variables for X_θ and Y_θ.

As the numerical values in Table 1 illustrate, a single-run estimator may yield a significant higher variance than a phantom estimator. This effect is caused for IPA from the fact that the derivative of the performance function itself comes into play and for SF estimator it is caused from the fact that likelihood ratios have notoriously high variance.

Table 1 Variance of the IPA, SF and phantom estimators

p	IPA	SF	Phantom	Coupled phantom
1	0.00160	0.02080	0.00480	0.00160
2	0.00512	0.00665	0.03123	0.00307
3	0.01575	0.01317	0.07842	0.00746
4	0.06511	0.03975	0.30045	0.02548
5	0.37011	0.17618	1.64838	0.12238

No type of estimator dominates the other. Single-run estimators are typically easy to implement but, at least in case of SF, usually tend to suffer from significant variance. A phantom estimator consumes computer storage for keeping track of the parallel phantoms, which makes the estimator usually cumbersome to implement, but has typically a low variance. As a rule of thumb, a single-run estimator should be applied whenever possible. However, counterexamples to this rule—besides the simple example given above—can be found in the literature as well. For instance, Pflug discusses in Sect. 4.3.2 of [8] a Markov chain example for which a phantom estimator outperforms a single-run estimator. Heidergott and Vázquez-Abad present in [11] a public transportation problem where a phantom estimator has considerable less variance than the single-run estimator. A thorough analysis of the relationship between single-run estimators and phantom type ones is a challenging subject for further research.

To conclude the section, we consider the scenario where the "phantom" system can be simulated via a change of measure, using the nominal system, thus implementing a weak derivative in a single-run experiment. This scenario explains the relationship between WD and SF. Consider a family $(P_{\theta,i} : 1 \leq i \leq n)$ of $\mathcal{D}$-differentiable Markov kernels on $(S, \mathcal{S})$ such that $P'_{\theta,i}$ is absolutely continuous with respect to $P_{\theta,i}$, in symbols, $P'_{\theta,i} \ll P_{\theta,i}$ for any $\theta \in \Theta$ and $1 \leq i \leq n$. Under uniform integrability conditions, using the same arguments as in Lemma 4.4 it can be shown that

$$
\begin{aligned}
&\frac{d}{d\theta}\mathbb{E}[g(X_\theta(n), \ldots, X_\theta(1))|X_\theta(0) = s_0] \\
&= \frac{d}{d\theta} \int g(s_n, \ldots, s_1) \prod_{i=1}^{n} P_{\theta,i}(ds_i; s_{i-1}) \\
&= \sum_{j=1}^{n} \iiint g(s_n, \ldots, s_1) \prod_{i=j+1}^{n} P_{\theta,i}(ds_i; s_{i-1}) P'_{\theta,j}(ds_j; s_{j-1}) \prod_{i=1}^{j-1} P_{\theta,i}(ds_i; s_{i-1}) \\
&= \sum_{j=1}^{n} \iiint g(s_n, \ldots, s_1) \frac{dP'_{\theta,j}}{dP_{\theta,j}}(s_j, s_{j-1}) \prod_{i=1}^{n} P_{\theta,i}(ds_i; s_{i-1}) \\
&= \mathbb{E}\left[g(X_\theta(n), \ldots, X_\theta(1)) \sum_{j=1}^{n} \frac{dP'_{\theta,j}}{dP_{\theta,j}}(X_\theta(j); X_\theta(j-1)) \,\middle|\, X_\theta(0) = s_0 \right],
\end{aligned}
$$

with $s_0 \in S$. Note that

$$\sum_{j=1}^{n} \frac{dP'_{\theta,j}}{dP_{\theta,j}}(\cdot;\cdot) = \sum_{j=1}^{n} \frac{d}{d\theta} \ln\left(P_{\theta,j}(\cdot;\cdot)\right),$$

which recovers the estimator called the score function, see Sect. 2.3.

It is worth noting that a single-run estimator can also be constructed if $P'_{\theta,i}$ fails to be dominated by P_θ. In such a case, one can take

$$Q_\theta = (1/3)P^+_{\theta,i} + (1/3)P^-_{\theta,i} + (1/3)P_{\theta,i}$$

as Markov kernel. Then, $P^{\pm}_{\theta,i}$, $P_{\theta,i} \ll Q_\theta$ and single-run estimator of the above type can be found, see [12] for details. However, manipulating the underlying Markov kernel in the above way is not always feasible and increases the variance of the estimator.

6 Discussion and Further Research

Building an estimator from the measure-valued differentiation formulas can be performed in a number of ways, depending on the implementation chosen. The estimator should (1) be easy to implement, (2 have low variance, and (3) have a low computational effort. Item (1) is often a matter of taste, while the two remaining criteria determine the *efficiency* of an estimator in simulation, and are often problem dependent.

Two measures μ, ν on $(S, \mathcal{S})$ are *orthogonal* if $A \in \mathcal{S}$ exists, such that $\mu(A) = 0$ and $\nu(A^c) = 0$; in symbols $\mu \perp \nu$. Applying the Hahn-Jordan decomposition for the $\mathcal{D}$-derivative of each of the one-step transition kernels P'_θ, the resulting measures are orthogonal: $P^+(\cdot; s) \perp P^-(\cdot; s)$ for all $s \in S$. As numerical examples show (see Chap. 4 in [8]), there is no guarantee that an orthogonal representation is always the one with the smallest variance.

Apart from the particular decomposition, $\mathcal{D}$-derivatives offer a further "degree of freedom". A $\mathcal{D}$-derivative only describes the marginal distribution of $(X^+_{\theta,j}(i), X^-_{\theta,j}(i))$ but not the *joint* distribution. A particular implementation of the estimation is to use the same underlying random variables to drive the evolution of each of the pairs $\{X^{\pm}_{\theta,j}(i), i = 1, 2, \ldots\}$, thus making these adapted to the natural filtration. Use of common random numbers for these processes further simplifies the estimation into

$$\begin{aligned}
&\frac{d}{d\theta}\mathbb{E}[g(X_\theta(n), \ldots, X_\theta(1))] \\
&\quad = \sum_{j=1}^{n} \mathbb{E}[c_{P_\theta}(X_\theta(j-1))(g(X^+_{\theta,j}(n), \ldots, X^+_{\theta,j}(1)) - g(X^-_{\theta,j}(n), \ldots, X^-_{\theta,j}(1)))].
\end{aligned}$$

Coupling via common random numbers is not necessarily optimal (in terms of variance reduction) for every performance function g. See [12] for detailed discussion on common random numbers for Markov chains. Nonetheless, examples abound where the "difference processes" $g(X^+_{\theta,j}) - g(X^-_{\theta,j})$ can be calculated recursively. The resulting estimators can have extremely low computational overhead, thus rendering very efficient estimation.

Lastly, it is not obvious when to choose a single-run estimator and when to implement a phantom estimator. For a given gradient estimation problem it generally depends on the particular problem which type of estimator is more efficient in terms of computation time. A thorough analysis of the trade-off between the two types of estimators in topic of further research.

References

1. Glasserman, P.: Gradient Estimation via Perturbation Analysis. Kluwer Academic, Boston (1991)
2. Fu, M., Hu, J.Q.: Conditional Monte Carlo. Kluwer Academic, Boston (1997)
3. Rubinstein, R., Shapiro, A.: Discrete Event Systems: Sensitivity Analysis and Optimization by the Score Function Method. Wiley, New York (1993)
4. Heidergott, B., Vázquez-Abad, F.J.: Measure-valued differentiation for stochastic processes: the random horizon case. Markov Process. Relat. Fields **12**, 509–536 (2006)
5. Heidergott, B., Vázquez-Abad, F.J.: Internet supplement, http://staff.feweb.vu.nl/bheidergott/JOTAsupplement.pdf
6. Cao, X.R.: Realization Probabilities: The Dynamics of Queueing Networks. Springer, New York (1994)
7. Ho, Y.C., Cao, X.R.: Perturbation Analysis of Discrete Event Dynamic Systems. Kluwer Academic, Boston (1991)
8. Pflug, G.: Optimization of Stochastic Models. Kluwer Academic, Boston (1996)
9. Dellacherie, C., Meyer, P.: Probabilities and Potential. North-Holland, Amsterdam (1978)
10. Heidergott, B., Hordjik, A., Weisshaupt, H.: Derivatives of Markov kernels and their Jordan decomposition. J. Appl. Anal. (2008, accepted for publication)
11. Heidergott, B., Hordjik, A.: Single-run gradient estimation via measure-valued differentiation. IEEE Trans. Autom. Control **49**, 1843–1846 (2004)
12. Glaserman, P., Yao, D.: Some guidelines and guarantees of common random numbers. Manag. Sci. **38**, 884–908 (1992)